

[0318] What is claimed is:

- 1 1. A computer-implemented method of discovering relationships between
- 2 items, comprising:
 - 3 accepting item selections from a plurality of users;
 - 4 generating a log for each user, each log containing identifiers for the user's
 - 5 item selections;
 - 6 accepting a query including at least one query item identifier;
 - 7 scoring the user logs, responsive to a degree of occurrence of the at least
 - 8 one query item identifier in the user logs, to generate user log
 - 9 scores; and
 - 10 determining at least one result item, responsive to a degree of occurrence
 - 11 in at least a subset of the scored user logs.
- 1 2. The computer-implemented method of claim 1, wherein a significance
- 2 of the occurrence is determined by a log likelihood ratio analysis and the deter-
- 3 mined result is responsive to the determined significance.
- 1 3. The computer-implemented method of claim 1, wherein a significance
- 2 of the occurrence is determined by a substantial equivalent of a log likelihood
- 3 ratio analysis and the determined result is responsive to the determined signifi-
- 4 cance.

1 4. The computer-implemented method of claim 1, wherein each item is a
2 video track and wherein accepting item selections comprises determining which
3 tracks are selected for playback.

1 5. The computer-implemented method of claim 1, wherein each item is a
2 music track and wherein accepting item selections comprises determining which
3 tracks are selected for playback.

1 6. The computer-implemented method of claim 5, further comprising:
2 generating a track list containing an identifier for each determined result
3 item comprising a music track.

1 7. The computer-implemented method of claim 6, further comprising:
2 deleting from the track list at least one identifier corresponding to a music
3 track already selected by the user.

1 8. The computer-implemented method of claim 6, further comprising:
2 playing the music tracks specified by the generated track list.

1 9. The computer-implemented method of claim 5, further comprising:
2 accepting a format schedule specifying music track categories for time pe-
3 riods; and

4 generating a track list conforming to the format schedule and containing
5 an identifier for each determined result item comprising a music
6 track.

1 10. The computer-implemented method of claim 5, wherein scoring the
2 user logs comprises determining a degree of occurrence in each user log of at
3 least one music track identified by the query item identifier.

1 11. The computer-implemented method of claim 5, wherein scoring the
2 user logs comprises determining a degree of occurrence in each user log of at
3 least one music track associated with an artist identified by the query item identi-
4 fier.

1 12. The computer-implemented method of claim 1, wherein accepting item
2 selections comprises receiving input provided by a user via a web page.

1 13. The computer-implemented method of claim 1, wherein accepting item
2 selections comprises receiving input specifying an item purchase by a user.

1 14. The computer-implemented method of claim 1, further comprising,
2 prior to determining the at least one result item, defining the subset of the scored
3 user logs responsive to the user log scores.

1 15. The computer-implemented method of claim 1, further comprising:
2 monitoring user behavior with respect to the selected items; and
3 adjusting the user log responsive to the monitored user behavior.

1 16. The computer-implemented method of claim 15, wherein monitoring
2 user behavior comprises at least one selected from the group consisting of:
3 detecting user input requesting that a selected item be repeated;
4 detecting user input requesting that a selected item be skipped;
5 detecting user input specifying a volume change; and
6 detecting user input specifying that a selected item be muted.

1 17. The computer-implemented method of claim 1, wherein accepting item
2 selections comprises receiving input provided by a user via an application for
3 playing tracks.

1 18. The computer-implemented method of claim 1, wherein accepting a
2 query comprises receiving a user log containing identifiers for a user's item selec-
3 tions.

1 19. The computer-implemented method of claim 1, wherein accepting a
2 query comprises receiving a first search term, the method further comprising:

3 generating a second search term containing an identifier for each deter-
4 mined result item.

1 20. The computer-implemented method of claim 19, further comprising at
2 least one of:

3 providing the second search term as input for a search engine; and
4 adding the second search term to a searchable portion of a document as-
5 sociated with the first search term.

1 21. The computer-implemented method of claim 1, further comprising:
2 periodically uploading the generated log.

1 22. The computer-implemented method of claim 1, further comprising:
2 outputting an advertisement relating to the determined at least one result
3 item.

1 23. The computer-implemented method of claim 22, wherein outputting
2 an advertisement comprises displaying at least one selected from the group con-
3 sisting of:

4 a web page;
5 a banner;
6 a portion of a web page; and
7 an animation.

1 24. The computer-implemented method of claim 1, further comprising:
2 outputting a notification relating to the determined at least one result
3 item.

1 25. The computer-implemented method of claim 24, wherein outputting a
2 notification comprises displaying a web page.

1 26. The computer-implemented method of claim 24, wherein outputting a
2 notification comprises sending a communication to a user.

1 27. The computer-implemented method of claim 26, wherein sending a
2 communication to a user comprises at least one selected from the group consist-
3 ing of:

4 transmitting an electronic mail message to the user;
5 telephoning the user; and
6 sending a direct mail item to the user.

1 28. The computer-implemented method of claim 1, wherein the deter-
2 mined result is responsive to a significance of the occurrence of the item in at
3 least a subset of the scored user logs, and wherein the significance is determined
4 by a log likelihood ratio analysis submethod comprising:
5 determining a total number of user logs N;

6 determining a number of user logs N_1 in a subset of user logs;
7 determining a number of user logs N_2 not in the subset of user logs;
8 determining a number of user logs k_{11} in the subset that include the item;
9 determining a number of user logs k_{12} not in the subset that include the
10 item;
11 determining a number of user logs $k_{21} = N_1 - k_{11}$ in the subset that do not
12 include the item;
13 determining a number of user logs $k_{22} = N_2 - k_{12}$ not in the subset that do
14 not include the item;
15 and determining a log likelihood ratio for the item.

1 29. The computer-implemented method of claim 28, wherein the log like-
2 lihood ratio is defined as:

3
$$\sum k_{ij} \log \frac{\pi_{ij}}{\mu_j}$$

4 where: $\pi_{ij} = \frac{k_{ij}}{N_j}$, $\mu_j = \sum_i \frac{k_{ij}}{N}$.

1 30. The computer-implemented method of claim 29, further comprising:
2 adjusting at least one of the k_{ij} values responsive to at least one selected
3 from the group consisting of:
4 the number of occurrences of the item in a user log;

- 5 the logarithm of the number of occurrences of the item in a user
- 6 log;
- 7 the number of occurrences of the item in all user logs;
- 8 the logarithm of the total number of users divided by the number
- 9 of users who have selected the item; and
- 10 a normalizing factor.

1 31. The computer-implemented method of claim 30, wherein the normal-
2 izing factor is $\frac{1}{\sqrt{\sum (S_j W_{ij})^2}}$, where S_j is a weight based on the number of occur-
3 rences of the item in all user logs and W_{ij} is a weight based on the number of oc-
4 currences of the item in a particular user log.

1 32. The computer-implemented method of claim 1, further comprising:
2 deleting from the determined at least one result item any result items al-
3 ready selected by a user associated with the query.

1 33. The computer-implemented method of claim 1, further comprising:
2 ranking the at least one result item responsive to the degree of signifi-
3 cance.

1 34. A computer-implemented method of discovering a relationship be-
2 tween a first item and a second item, comprising:

3 determining a total number of item groups N;
4 determining a number of item groups N_1 in a subset of item groups, the
5 subset of item groups being defined as including those item
6 groups that contain a second item;
7 determining a number of item groups N_2 not in the subset of item groups;
8 determining a number of item groups k_{11} in the subset that contain the
9 first item;
10 determining a number of item groups k_{12} not in the subset that contain the
11 first item;
12 determining a number of item groups $k_{21} = N_1 - k_{11}$ in the subset that do
13 not contain the first item;
14 determining a number of item groups $k_{22} = N_2 - k_{12}$ not in the subset that
15 do not contain the first item;
16 and determining a log likelihood ratio.

1 35. The computer-implemented method of claim 34, wherein the log like-
2 lihood ratio is defined as:

$$3 \quad \sum k_{ij} \log \frac{\pi_{ij}}{\mu_j}$$

4 where: $\pi_{ij} = \frac{k_{ij}}{N_j}$, $\mu_j = \sum_i \frac{k_{ij}}{N}$.

1 36. The computer-implemented method of claim 35, wherein each item
2 group comprises a document.

1 37. The computer-implemented method of claim 35, further comprising:
2 adjusting at least one of the k_{ij} values responsive to at least one selected
3 from the group consisting of:
4 the number of occurrences of the item in a document;
5 the logarithm of the number of occurrences of the item in a docu-
6 ment;
7 the number of occurrences of the item in all documents;
8 the logarithm of the total number of documents divided by the
9 number of documents that include the item; and
10 a normalizing factor.

1 38. The computer-implemented method of claim 37, wherein the normal-
2 izing factor is $\frac{1}{\sqrt{\sum (S_j W_{ij})^2}}$, where S_j represents the number of occurrences of
3 the item in all documents and W_{ij} represents the number of occurrences of the
4 item in a particular document.

1 39. A system for discovering relationships among items, comprising:
2 a user interface for accepting item selections from a plurality of users;

3 at least one log database, coupled to the user interface, for storing a log for
4 each user, each log containing identifiers for the user's item
5 selections;
6 a query input device for accepting a query including at least one query
7 item identifier; and
8 a recommendation engine, coupled to the log database and to the query
9 input device, for scoring the user logs, responsive to a degree of
10 occurrence, to generate user log scores, and for determining at
11 least one result item, responsive to a degree of occurrence in at
12 least a subset of the scored user logs.

1 40. The system of claim 39, wherein the significance of the occurrence is
2 determined by a log likelihood ratio analysis and the recommendation engine
3 determines the at least one result item responsive to the determined significance.

1 41. The system of claim 39, wherein the significance of the occurrence is
2 determined by a substantial equivalent of a log likelihood ratio analysis and
3 wherein the recommendation engine determines the at least one result item re-
4 sponsive to the determined significance.

1 42. The system of claim 39, wherein each item is a video track and wherein
2 the user interface accepts item selections by determining which tracks are se-
3 lected for playback.

1 43. The system of claim 39, wherein the user interface accepts item selec-
2 tions by determining which tracks are selected for purchase.

1 44. The system of claim 39, wherein each item is a music track and
2 wherein the user interface accepts item selections by determining which tracks
3 are selected for playback.

1 45. The system of claim 44, wherein the user interface comprises an online
2 jukebox.

1 46. The system of claim 45, wherein the online jukebox monitors user be-
2 havior with respect to the selected items and adjusts the user log scores respon-
3 sive to the monitored user behavior.

1 47. The system of claim 46, wherein the online jukebox monitors user be-
2 havior by detecting at least one selected from the group consisting of:
3 user input requesting that a selected item be repeated; and
4 user input requesting that a selected item be skipped; and

5 user input specifying a volume change; and
6 user input specifying that a selected item be muted.

1 48. The system of claim 47, further comprising:
2 a track list generator, coupled to the recommendation engine, for generat-
3 ing a track list containing an identifier for each determined re-
4 sult item comprising a music track.

1 49. The system of claim 44, further comprising:
2 a music player, coupled to the track list generator, for playing the music
3 tracks specified by the generated track list.

1 50. The system of claim 44, further comprising:
2 a format scheduler, for accepting a format schedule specifying music track
3 categories for time periods; and
4 a track list generator, coupled to the recommendation engine and to the
5 format scheduler, for generating a track list conforming to the
6 format schedule and containing an identifier for each deter-
7 mined result item comprising a music track.

1 51. The system of claim 39, wherein the query input device receives a user
2 log containing identifiers for a user's item selections.

1 52. The system of claim 39, wherein the query input device receives a first
2 search term, the system further comprising:
3 a search term generator, coupled to the recommendation engine, for gen-
4 erating a second search term containing an identifier for each
5 determined result item and for providing the second search
6 term as input for a search engine.

1 53. The system of claim 39, wherein the query input device receives a first
2 search term, the system further comprising:
3 a search term generator, coupled to the recommendation engine, for gen-
4 erating a second search term containing an identifier for each
5 determined result item and for providing the second search
6 term to be added to a searchable portion of a document associ-
7 ated with the first search term.

1 54. The system of claim 39, further comprising:
2 an advertisement output device, coupled to the recommendation engine,
3 for outputting an advertisement relating to the determined at
4 least one result item.

1 55. The system of claim 54, wherein the advertisement output device dis-
2 plays at least one selected from the group consisting of:

3 a web page;
4 a banner;
5 a portion of a web page; and
6 an animation.

1 56. The system of claim 39, further comprising:
2 a notification output, coupled to the recommendation engine, for output-
3 ting a notification relating to the determined at least one result
4 item.

1 57. The system of claim 56, wherein the notification output device displays
2 at least one selected from the group consisting of:
3 a web page;
4 a banner;
5 a portion of a web page; and
6 an animation.

1 58. The system of claim 56, wherein the notification output device sends a
2 communication to a user.

1 59. A computer-readable medium comprising computer-readable code for
2 discovering relationships between items, comprising:

3 computer-readable code adapted to accept item selections from a plurality
4 of users;
5 computer-readable code adapted to generate a log for each user, each log
6 containing identifiers for the user's item selections;
7 computer-readable code adapted to accept a query including at least one
8 query item identifier;
9 computer-readable code adapted to score the user logs, responsive to a
10 degree of occurrence of the at least one query item identifier in
11 the user logs, to generate user log scores; and
12 computer-readable code adapted to determine at least one result item, re-
13 sponsive to a degree of occurrence in at least a subset of the
14 scored user logs.

1 60. The computer-readable medium of claim 59, wherein a significance of
2 the occurrence is determined by a log likelihood ratio analysis and the deter-
3 mined result is responsive to the determined significance.

1 61. The computer-readable medium of claim 59, wherein a significance of
2 the occurrence is determined by a substantial equivalent of a log likelihood ratio
3 analysis and the determined result is responsive to the determined significance.

1 62. The computer-readable medium of claim 59, wherein each item is a
2 video track and wherein the computer-readable code adapted to accept item se-
3 lections comprises computer-readable code adapted to determine which tracks
4 are selected for playback.

1 63. The computer-readable medium of claim 59, wherein each item is a
2 music track and wherein the computer-readable code adapted to accept item se-
3 lections comprises computer-readable code adapted to determine which tracks
4 are selected for playback.

1 64. The computer-readable medium of claim 63, further comprising:
2 computer-readable code adapted to generate a track list containing an
3 identifier for each determined result item comprising a music
4 track.

1 65. The computer-readable medium of claim 64, further comprising:
2 computer-readable code adapted to delete from the track list at least one
3 identifier corresponding to a music track already selected by the
4 user.

1 66. The computer-readable medium of claim 64, further comprising:

2 computer-readable code adapted to play the music tracks specified by the
3 generated track list.

1 67. The computer-readable medium of claim 63, further comprising:
2 computer-readable code adapted to accept a format schedule specifying
3 music track categories for time periods; and
4 computer-readable code adapted to generate a track list conforming to the
5 format schedule and containing an identifier for each deter-
6 mined result item comprising a music track.

1 68. The computer-readable medium of claim 63, wherein the computer-
2 readable code adapted to score the user logs comprises computer-readable code
3 adapted to determine a degree of occurrence in each user log of at least one mu-
4 sic track identified by the query item identifier.

1 69. The computer-readable medium of claim 63, wherein the computer-
2 readable code adapted to score the user logs comprises computer-readable code
3 adapted to determine a degree of occurrence in each user log of at least one mu-
4 sic track associated with an artist identified by the query item identifier.

1 70. The computer-readable medium of claim 59, wherein the computer-
2 readable code adapted to accept item selections comprises computer-readable
3 code adapted to receive input provided by a user via a web page.

1 71. The computer-readable medium of claim 59, wherein the computer-
2 readable code adapted to accept item selections comprises computer-readable
3 code adapted to receive input specifying an item purchase by a user.

1 72. The computer-readable medium of claim 59, further comprising, com-
2 puter-readable code adapted to, prior to determine the at least one result item,
3 define the subset of the scored user logs responsive to the user log scores.

1 73. The computer-readable medium of claim 59, further comprising:
2 computer-readable code adapted to monitor user behavior with respect to
3 the selected items; and
4 computer-readable code adapted to adjust the user log scores responsive
5 to the monitored user behavior.

1 74. The computer-readable medium of claim 73, wherein the computer-
2 readable code adapted to monitor user behavior comprises at least one selected
3 from the group consisting of:
4 computer-readable code adapted to detect user input requesting that a se-
5 lected item be repeated;
6 computer-readable code adapted to detect user input requesting that a se-
7 lected item be skipped;

8 computer-readable code adapted to detect user input specifying a volume
9 change; and
10 computer-readable code adapted to detect user input specifying that a se-
11 lected item be muted.

1 75. The computer-readable medium of claim 59, wherein the computer-
2 readable code adapted to accept item selections comprises computer-readable
3 code adapted to receive input provided by a user via an application for playing
4 tracks.

1 76. The computer-readable medium of claim 59, wherein the computer-
2 readable code adapted to accept a query comprises computer-readable code
3 adapted to receive a user log containing identifiers for a user's item selections.

1 77. The computer-readable medium of claim 59, wherein the computer-
2 readable code adapted to accept a query comprises computer-readable code
3 adapted to receive a first search term, the computer-readable medium further
4 comprising:
5 computer-readable code adapted to generate a second search term con-
6 taining an identifier for each determined result item.

1 78. The computer-readable medium of claim 77, further comprising at
2 least one of:

3 computer-readable code adapted to provide the second search term as in-
4 put for a search engine; and
5 computer-readable code adapted to add the second search term to a
6 searchable portion of a document associated with the first
7 search term.

1 79. The computer-readable medium of claim 59, further comprising:
2 computer-readable code adapted to periodically upload the generated log.

1 80. The computer-readable medium of claim 59, further comprising:
2 computer-readable code adapted to output an advertisement relating to
3 the determined at least one result item.

1 81. The computer-readable medium of claim 80, wherein the computer-
2 readable code adapted to output an advertisement comprises computer-readable
3 code adapted to display at least one selected from the group consisting of:
4 a web page;
5 a banner;
6 a portion of a web page; and
7 an animation.

1 82. The computer-readable medium of claim 59, further comprising:

2 computer-readable code adapted to output a notification relating to the
3 determined at least one result item.

1 83. The computer-readable medium of claim 82, wherein the computer-
2 readable code adapted to output a notification comprises computer-readable
3 code adapted to display a web page.

1 84. The computer-readable medium of claim 82, wherein the computer-
2 readable code adapted to output a notification comprises computer-readable
3 code adapted to send a communication to a user.

1 85. The computer-readable medium of claim 84, wherein the computer-
2 readable code adapted to send a communication to a user comprises at least one
3 selected from the group consisting of:
4 computer-readable code adapted to transmit an electronic mail message to
5 the user;
6 computer-readable code adapted to telephone the user; and
7 computer-readable code adapted to send a direct mail item to the user.

1 86. The computer-readable medium of claim 59, wherein the determined
2 result is responsive to a significance of the occurrence of the item in at least a
3 subset of the scored user logs, and wherein the computer-readable code adapted
4 to determine a binomial log likelihood ratio for an item comprises computer-

5 readable code adapted to determine the result by a log likelihood ratio analysis
6 submethod.

1 87. The computer-readable medium of claim 86, wherein the computer-
2 readable code adapted to determine the result by a log likelihood ratio analysis
3 submethod comprises:

4 computer-readable code adapted to determine a total number of users N ;
5 computer-readable code adapted to determine a number of users N_1 in a
6 subset of users;
7 computer-readable code adapted to determine a number of users N_2 not in
8 the subset of users;
9 computer-readable code adapted to determine a number of users k_{11} in the
10 subset that selected the item;
11 computer-readable code adapted to determine a number of users k_{12} not
12 in the subset that selected the item;
13 computer-readable code adapted to determine a number of users $k_{21} = N_1$
14 - k_{11} in the subset that did not select the item;
15 computer-readable code adapted to determine a number of users $k_{22} = N_2$
16 - k_{12} not in the subset that did not select the item; and
17 computer-readable code adapted to determine a log likelihood ratio for
18 the item.

1 88. The computer-readable medium of claim 87, wherein the log likeli-
2 hood ratio is defined as:

3 $\sum k_{ij} \log \frac{\pi_{ij}}{\mu_j}$

4 where: $\pi_{ij} = \frac{k_{ij}}{N_j}$, $\mu_j = \sum_i \frac{k_{ij}}{N}$.

1 89. The computer-readable medium of claim 59, wherein the computer-
2 readable code adapted to determine the result by a log likelihood ratio analysis
3 submethod further comprises:

4 computer-readable code adapted to adjust at least one of the n_{ij} values re-
5 sponsive to at least one selected from the group consisting of:
6 the number of occurrences of the item in a user log;
7 the logarithm of the number of occurrences of the item in a user
8 log;
9 the number of occurrences of the item in all user logs;
10 the logarithm of the total number of users divided by the number
11 of users who have selected the item; and
12 a normalizing factor.

1 90. The computer-readable medium of claim 89, wherein the normalizing
2 factor is $\frac{1}{\sqrt{\sum(S_jW_{ij})^2}}$, where S_j is a weight based on the number of occurrences
3 of the item in all user logs and W_{ij} is a weight based on the number of occur-
4 rences of the item in a particular user log.

1 91. The computer-readable medium of claim 59, further comprising:
2 computer-readable code adapted to delete from the determined at least
3 one result item any result items already selected by a user asso-
4 ciated with the query.

1 92. The computer-readable medium of claim 59, further comprising:
2 computer-readable code adapted to rank the at least one result item re-
3 sponsive to the degree of significance.

1 93. A computer-readable medium comprising computer-readable code for
2 discovering a relationship between a first item and a second item, comprising:
3 computer-readable code adapted to determine a total number of item
4 groups N;
5 computer-readable code adapted to determine a number of item groups
6 N_1 in a subset of item groups, the subset of item groups being

7 defined as including those item groups that contain a second
8 item;
9 computer-readable code adapted to determine a number of item groups
10 N₂ not in the subset of item groups;
11 computer-readable code adapted to determine a number of item groups
12 k₁₁ in the subset that contain the first item;
13 computer-readable code adapted to determine a number of item groups
14 k₁₂ not in the subset that contain the first item;
15 computer-readable code adapted to determine a number of item groups
16 k₂₁ = N₁ - k₁₁ in the subset that do not contain the first item;
17 computer-readable code adapted to determine a number of item groups
18 k₂₂ = N₂ - k₁₂ not in the subset that do not contain the first item;
19 and
20 computer-readable code adapted to determine a log likelihood ratio.

1 94. The computer-readable medium of claim 93, wherein the log likeli-
2 hood ratio is defined as:

3
$$\sum k_{ij} \log \frac{\pi_{ij}}{\mu_j}$$

4 where: $\pi_{ij} = \frac{k_{ij}}{N_j}$, $\mu_j = \sum_i \frac{k_{ij}}{N}$.

1 95. The computer-readable medium of claim 93, wherein each item group
2 comprises a document.

1 96. The computer-readable medium of claim 93, further comprising:
2 computer-readable code adapted to adjust at least one of the k_{ij} values re-
3 sponsive to at least one selected from the group consisting of:
4 the number of occurrences of the item in a document;
5 the logarithm of the number of occurrences of the item in a docu-
6 ment;
7 the number of occurrences of the item in all documents;
8 the logarithm of the total number of documents divided by the
9 number of documents that include the item; and
10 a normalizing factor.

1 97. The computer-readable medium of claim 96, wherein the normalizing
2 factor is $\frac{1}{\sqrt{\sum (S_j W_{ij})^2}}$, where S_j represents the number of occurrences of the item
3 in all documents and W_{ij} represents the number of occurrences of the item in a
4 particular document.